7000 years of noun phrase evolution

Freek Van de Velde
University of Leuven

Language change

- The rate of language change is:

  fast when compared to biological change, but

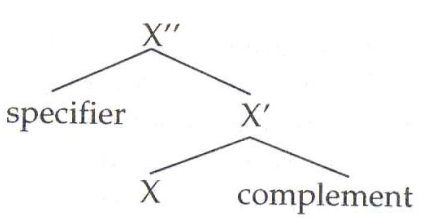
  slow when compared to other human cultural products (e.g. fashion):

  ⇨ Cultural evolution can be Lamarckian and guided, but less so in language

  - **Phonology**: sounds may change within one generation
  - **Lexicon**: words have a linguistic half-life of roughly 2k – 4k years, very frequent words up to 10k – 20k years (Pagel et al. 2013)
  - **Grammar** (syntax and morphology):
    - more stable than phonology
    - *Eppur Si Muove*
    - Long-term 'drifts'

Evolution in syntax?

- Received wisdom: 'universality'
  - Formal grammar: UG



  - Functional grammar: Greenberg

    "In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object."

  - All languages are equally complex: Boas, Sapir, Whorf, Chomsky …

    "It seems to me that there is a principle of conservation of complexity, under which any rise in complexity in one system in the language results in a decrease of complexity elsewhere." (Aya Katz, 2011/03/18 on FUNKNET)

    "A language which is simple and regular in one respect is likely to be complex and confusing in others. There seems to be a trading relationship between the different parts of the grammar which we do not fully understand." (Aitchison 1991:214)

    (more citations in Miestamo 2008)

Evolution in syntax?

1. The universality thesis has come under attack

- Grammatical categories are language-specific (Dryer 1997; Croft 2001; Haspelmath 2007; Evans & Levinson 2009).

2. Different categories ⇒ different degrees of complexity

- Some languages are more complex than others (Perkins 1992; McWhorther 2001; Kusters 2003; Dahl 2004; Everett 2005; Sampson 2009; Sampson et al. 2009; Trudgill 2010). But: it is not clear what 'complexity' exactly is.

3. Complexity: differences between languages ⇒ between language stages

- Languages can become more complex (Martinet 1973; Deutscher 2000; Heine & Kuteva 2007; Givón 2009; Van de Velde 2010) (or, alternatively, more simple, Kusters 2003)

Evolution in syntax?

- Whence the opposition against 3 ('change in linguistic complexity')?
  - Revives the 'organic' view on language change advocated by Bopp, Grimm, Humboldt, Schlegel, Schleicher en Max Müller (Morpurgo-Davies 1998:86-88, 196-200)
  - Revives romantic Haeckelian recapitulation theory
  - Flies in the face of the analyticity – syntheticity trade-off (on which more later)
  - Suggests that people used to be stupid: not capable of dealing with hypotaxis … (see Weerman 1986)
  - But: smaller ('primitive') languages are often more complex than big languages

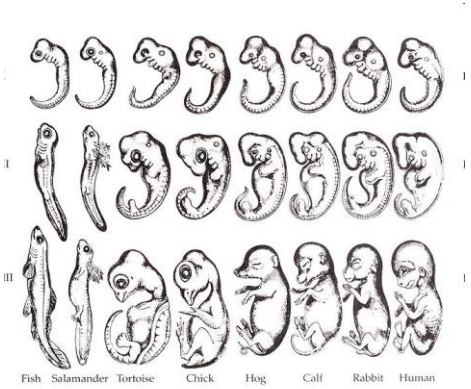


Illustration: Romanes 1892 (from Wikipedia)

Evolution in syntax?

- Taking one step back: What is complexity exactly?
  - Different opinions in the literature (Hawkins 2004; Dahl 2004; Miestamo et al. 2008; Sampson et al. 2009)
  - Hypothesis Lupyan & Dale (2010):

Morphologic complexity

- redundance (cf. Trudgill 2010:308)
- favouring L1 acquisition
- patterns with smaller languages
- low-contact
- 'esoteric languages'
- e.g. evidential morphology, demonstrative affixes

Analytic-syntactic complexity

- transparency (cf. Trudgill 2010:308)
- favouring L2 acquisition (cf. Kusters 2003:48ff., 2008)
- patterns with bigger languages
- high-contact
- exoteric languages
- e.g. evidential adverbs, articles

Evolution in syntax?

- Morphological complexity ('syntheticity'): inflection

  Quechua (cited in Hengeveld 1989:142)
  pay-ka *shamu-nga-m-ári*
  he-TOP come-FUT.3-FIRSTHAND-REINF
  'He will definitely come'

  English
  We *come*(-Ø)

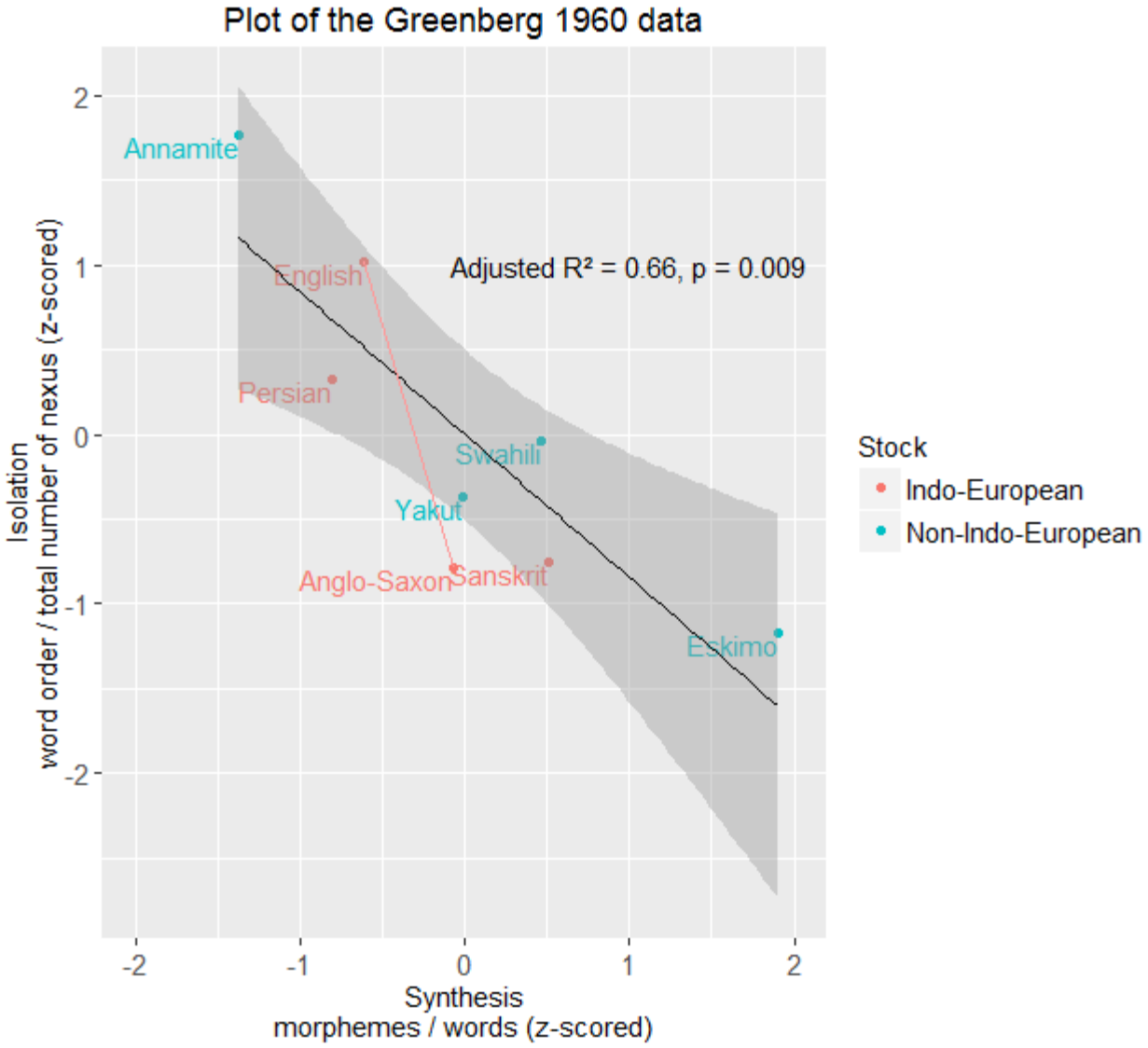
- Syntactic complexity ('analyticity'): word order and function words

  English
  John thought that Bill hates the fact that Mary loves Kate.
  all these beautiful paintings (no permutations possible)
  He could have worked

- Time-honoured idea: Trade-off between analyticity and syntheticity:

  "En Europe les langues dérivées du latin, et l'anglais, ont une grammaire tout analytique (...) synthétiques dans leur origine (...) elles penchent fortement vers les formes analytiques" (Von Schlegel 1846:161, cited in Szmrecsanyi 2012)

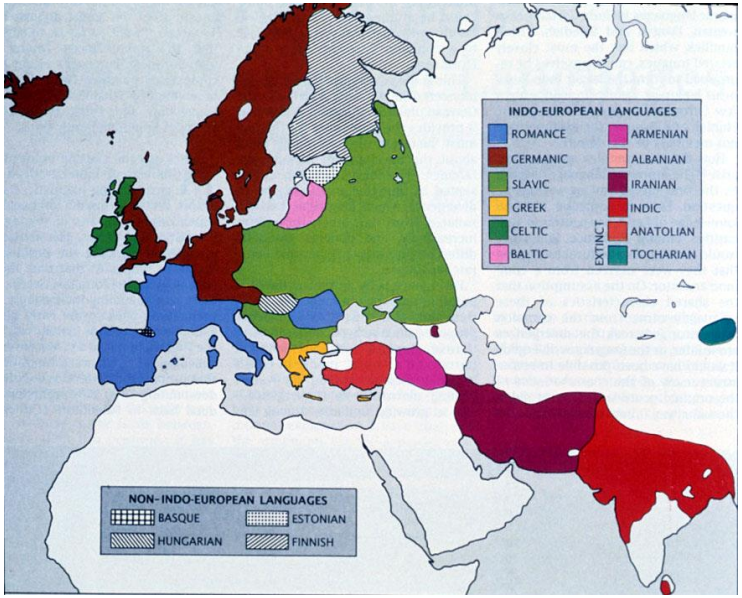
Trade-off between morphology and syntax

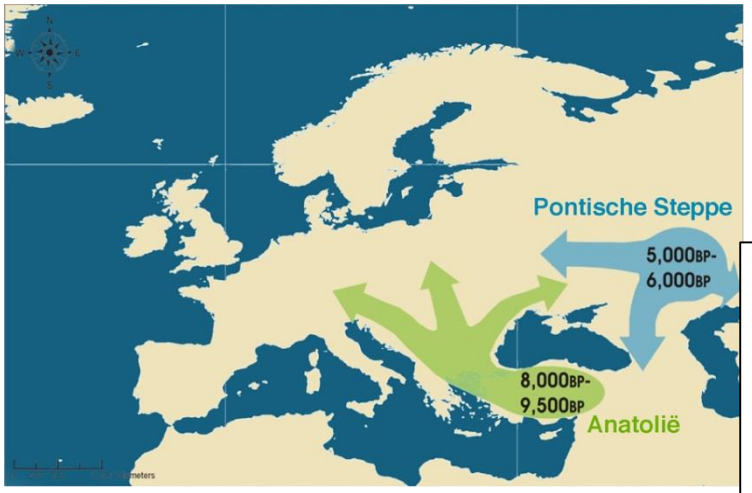


FROM PROTO-INDO-EUROPEAN TO PRESENT-DAY DUTCH AND ENGLISH

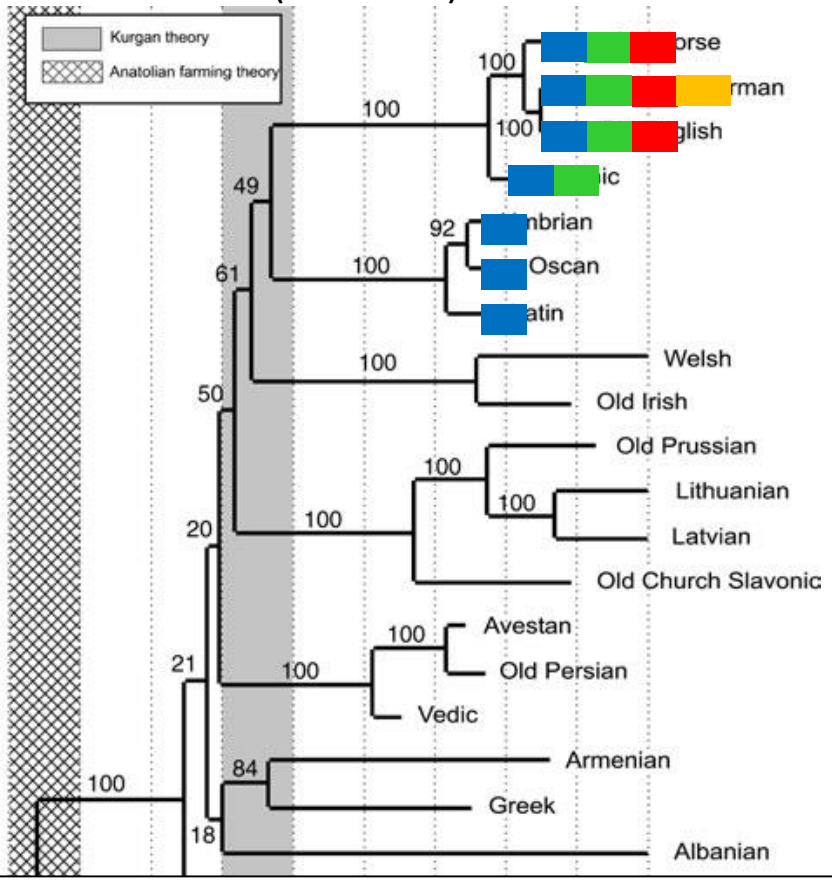
Some background

new tense from perfect
Grimm's law
umlaut
Grimm's second law





Atkinson et al. (2005:215)



Difficulties:

- lateral transfer (contact, wave model) (m/bh in dative in Germanic, Baltic and Slavic)
- homoplasy (parallel drift) (/o:/ > /u:/ in German and English)
- differential speed in language change

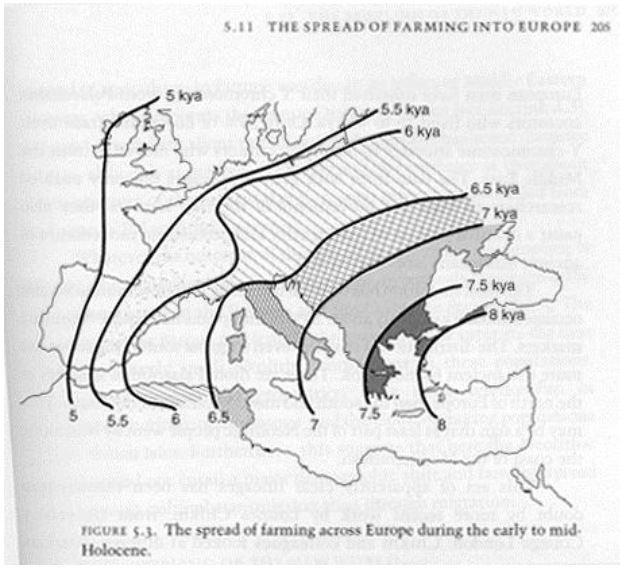
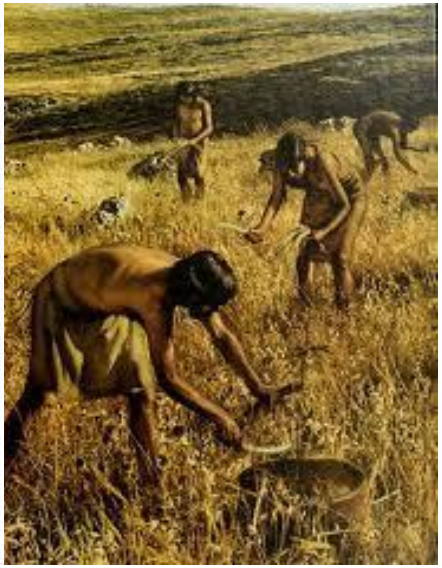
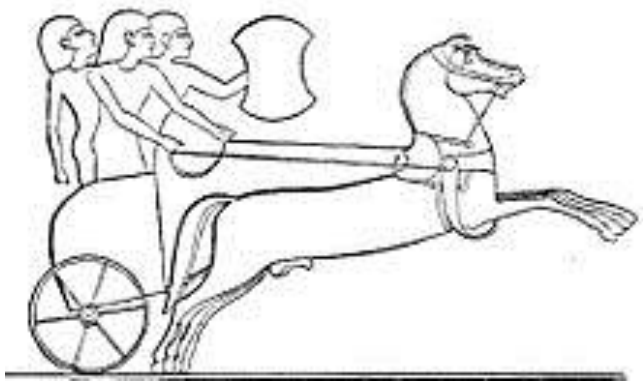


FIGURE 5.3. The spread of farming across Europe during the early to mid-Holocene.





The Anatolic hypothesis

Complexity drift in Indo-European

- Research in simplexification / complexification is often focused on morphology, i.e. the rise, maintenance or loss of nominal case and verbal inflection

- What about syntactic complexity?
  - Deutscher (2000): rise of complement clauses in Akkadian is related to the complexification of the Babylonic society.
  - Complex verbal clusters in the history of Dutch (Coupé 2015)
  - Complex noun phrases (NPs) are not universal, but are a typical feature of (Indo-)European language (Rijkhoff 1998, 2002).

- Indo-European languages show a drift of 'deflection'. Concomitantly (but not necessarily as an immediate mechanic result) they rely increasingly on hierarchical, integrated phrases.

Nouns and noun phrases

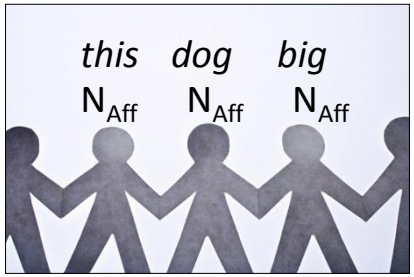
- A noun phrase is the syntactic group comprising the noun and its dependents.

- Simplifying: the NP is the group of words that can occupy:
  - the position of the direct object (*I saw* ___)
  - the position after a preposition (*through* ___)

- Do all languages and language stages have noun phrases? Non-universality of NP structure: see Matthews 1981, 2007; Rijkhoff 1998; Evans & Levinson 2009: 440-442

- How do languages differ in the expression of NP-like structures?

  – morphological expression of nominal categories (definiteness, number, …), no fixed order (non-configurational) ⇨ **flat ('democratic')**



    Cipa-yi *ṭuku-yu* *yaun-tu* *yani* *cayi*
    this-ERG dog-ERG big-ERG white_man bite
    'this big dog bit/bites the white man'
    (Kalkatungu, Blake 1983:145, quoted in Rijkhoff 2002:19-20)

  – specialised slots for the expression of nominal categories (definiteness, number ...) , fixed order (configurational) ⇨ **hierarchical ('feudal')**

    [$_{NP}$ *this*$_{D}$ [*big*$_{A}$ [*dog*$_{N}$]]]

  – Mixed:

    a sheyn meydl vs. *a meydl a sheyne*
    a pretty girl a girl, a petty (one)
    (Yiddish, Jacobs et al. 1994:408, quoted in Rijkhoff 1998:324)

Against universality of NP structure

- What are reliable diagnostics of phrasal structure?
  - **Semantics**: 'what belongs together forms a constituent'
    - Problem: What does 'belong to' mean? All sorts of element that 'belong to the noun' are cross-linguistically found to be expressed outside the NP, verbally or adverbially (quantifiers, deixis, modifiers, secondary predicates, external possessors ...) (see also Rijkhoff 2015)
    - E.g. *Ihm schmerzt der Bauch* vs. *His belly aches*
  - **Pronominalisation** (Zwicky 1978; McCawley 1988; Carnie 2002:50-51; Van Valin 2001:111-112)
    - *He ate the remaining slice of cranberry cake* ~ *He ate it*
    - Problem: does not work (Croft 2001:188; Keizer 2011, 2012; Van de Velde 2012)

Against universality of NP structure

- What are reliable diagnostics of phrasal structure?
  - **Contiguity**: constituency relies on contiguous expression (see for theoretical underpinnings: Robinson 1970; Koster 1978; Sturm 1986; Verhagen 1992, 1996; Langacker 1997; Evans & Levinson 2009:440-442; Van de Velde 2009a, 2012)
    - Problem: extraposition of postmodifiers (We have several important books in stock about global warming)
    - But if they can be extraposed, they are introduced by a 'relator', which is an alternative to syntactic phrase structure. The existence of phrase structure in a language does not preclude the use of other means to express semantic relations.
    - In general postmodifiers are less integrated (Verhaar 1998:102; Plank 2003:346)
  - **Ordering constraints** (in the sense of Payne & Huddleston 2002:452): subparts of constituents shouldn't be freely switched around.

What counts as an noun phrase?

- Constituency encoded in language by phrase structure
  1. Contiguity / inseparability
  2. Fixed relative position

- Other means of nexus are *not* consituency:
  1. Case marking
  2. Prepositions
  3. Agreement
  ...

No noun phrases in Proto-Indo-European

- PIE has a word-based syntax, rather than a phrase-based syntax

  "die Beziehung selbst bekommt keinen unmittelbaren Ausdruck, sie kann keinen bekommen; die Wörter werden zunächst nur aneinander gerückt, wie die Steine einer kyklopischen Mauer; kein Kitt oder Mörtel verbindet sie, die Kluft wird schliesslich durch kleine Keile ausgefüllt" (Schuchardt, 1922)

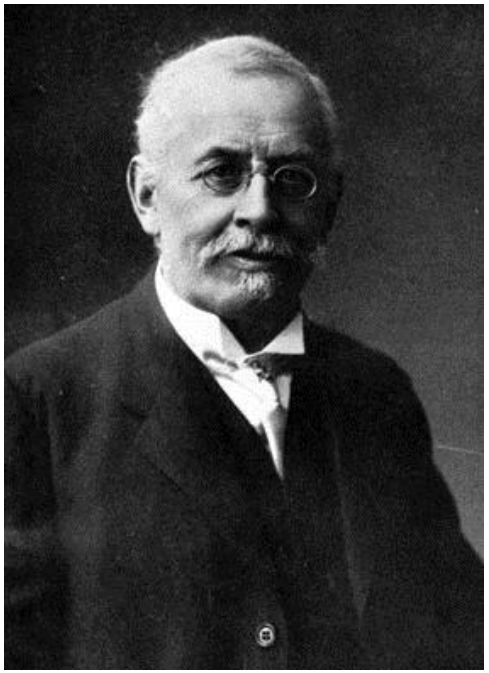
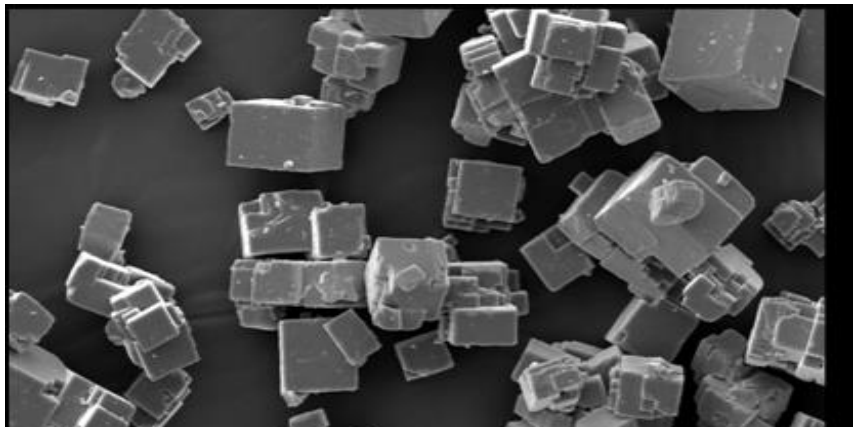


Illustration: Cyclop wall (Courtesy Joop van der Horst)

Crystallisation of NP structure from PGm to PdWG

- Hypothesis: IE saw the emergence and expansion of phrasal structure in the NP, continuing to this day in West-Germanic



http://www.tvt.cbi.uni-erlangen.de/eng/research/crystalisation/crystallisation.htm

Rise of noun phrase structure

- Controversial …
- … but not outragous, similar ideas in:
  - Meillet & Vendryes (1924)
  - Meillet (1934)
  - Von Wartburg (1962)
  - Krause (1968)
  - Lehmann (1974)
  - Faarlund (2001)
  - Van de Velde (2009a, b)
  - Luraghi (2010)
  - Perridon & Sleeman (2011)
  - Ledgeway (2011, 2012)
  - Van de Velde et al. (2014)
  - Börjars et al. (2015)

Some quotes

- "Progress continues to be made in our understanding of the ways in which grammatical structures can develop over time. Considerable attention has been focused on processes by which larger, looser patterns are crystallized into tighter, more compact ones. Frequently used discourse patterns, for example, such as particular constituent order, can become routinized and rigidified in syntactic structure." (Mithun 2008:68)

- Bybee (2002): "items that are used together fuse together”

- Luraghi (2010:220): "[C]onstituency had become relevant for NPs."

Rise of noun phrase structure

- PIE did not have NPs of the type [$_{NP}$ *black sheep*]
  - PIE does not make a distinction between nouns and adjectives
    - Alleged adjectives are inflectionally indistinguishable from nouns
    - Alleged nouns can be used as adjectives without derivational morphology
    - Adjective-noun agreement is often violated
    - Gradation is not restricted to adjectives
  - PIE does not have attributive modifiers
    - Alleged modifiers are not contiguously expressed
    - Alleged modifiers can have their own NP markers

Rise of noun phrase structure

- PIE does not make a distinction between nouns (*sheep*) and adjectives (*black*)
  - As in Quechua or Turkish (Schachter 1985; Hengeveld 1992; Stassen 1997)
  - Alleged adjectives are inflectionally indistinguishable from nouns (Bammesberger 1992:52; Kurzová 1993:41)
  - Alleged nouns can be used as adjectives without derivational morphology (and vice versa) (Brugmann & Delbrück 1889:436, 444-448)

Example	**Noun**	**Adjective**
nātus (Latin)	'son'	'born'
ūber (Latin)	'udder'	'rich'
srutá-m, *srutá-s* (Old Indic)	'flood'	'streaming'
pinut-ós, *pinut-ḗ* (Greek)	'reason'	'reasonable'
takšul (Hittite)	'peace'	'friendly'

Rise of noun phrase structure

- PIE does not make a disinction between nouns and adjectives
  - Adjective-noun agreement is often violated (Lehmann 1974:69-70; Brugmann & Delbrück 1893:402-410; Hirt 1927:330-332; 1934:143-144; Lehmann 1974:69-73; Van de Velde 2009a:171-173, 2009b:1026-1027)
  - Gradation is not restricted to adjectives (Brugmann and Delbrück 1893:415-416; Hirt 1934: 147; Lehmann 1995:156, 227)
    - *kavítara* (poet:COMPARATIVE) / *kavítama* (poet:SUPERLATIVE) (Old Indic)
    - *basileúteros* (king:COMPARATIVE) / *basileútatos* (koning:SUPERLATIVE) (Greek)

Rise of noun phrase structure

- PIE does not have attributive modifiers
  - Alleged modifiers are not contiguously expressed

fílōi	*epepeítheth‘*	***hetaírōi***
beloved:DAT.M.SG	be.persuaded:IND.IPFV.3SG	companion:DAT.M.SG

'he obeyed the beloved companion‘ (Greek, Hirt 1937:230)

meo	*tu*	*epistulam*	*dedisti*	***servo***
my:DAT.SG	you	letter:ACC.SG	gave	slave:DAT.SG

‘to my slave you gave a letter?’ (Latin, Perridon & Sleeman 2011: 4)

juvám	***śvētám***	*Pēdáva …*	*adattam*	***áśvam***
you:NOM.PL	white:ACC.M.SG	P.	give:ACT.IND.IPFV.2DU	horse:ACC.M.SG

'the white horse you gave to P.‘ (Old Indic, Hirt 1937:231)

hantezziyass*=a=*	*as*	***ÌR***	*lē*	
first:NOM	and	he:NOM	servant	not

'Let him not be one of [my] ministers‘ (Hittite, Luraghi 1990: 171)

Rise of noun phrase structure

- PIE does not have attributive modifiers
  - Alleged modifiers can have their own NP markers (Van de Velde 2009a:183-187)
    - CASE (~ dependency reversal, Malchukov 2000)

      ho *hḗmisus* *toû* *chrónou*
      art:NOM.M.SG half:NOM.M.SG art:GEN.M.SG time:GEN.M.SG
      'half of the time‘ (Greek, Kühner & Gerth 1963:279)

      epì *tēi* *hēmiseíai* *tēs* *gēs*
      on ART:DAT.F.SG half:DAT.F.SG ART:GEN.SG earth:GEN.F.SG
      'on half of the earth' (Greek, Kühner & Gerth 1963:279)

    - ARTICLES

      hoi *gérontes* *hoi* *palaioì*
      ART:NOM.M.PL old_men:NOM.PL ART:NOM.M.PL elderly:NOM.M.PL
      'the elderly old men' (Greek, Krüger & Cooper 2001:429)

Crystallisation of noun phrase structure

- Emergence of phrasal structure in Germanic (2000 BC – 500 AD)
- Gradual integration of erstwhile free modifiers in the NP
- Follows a motivated path:
  - Higher scope modifiers resist incorporation longer (adjectives > determiners > peripheral modifiers)
  - Marked modifiers resist incorporation longer than prototypical modifiers (adjectives > participles, quantifiers ...)

Period	**Structure**	**New slot**	**example**
PIE	N	Noun	*sheep*
PGm	$[_{NP}$ A [N]]	Slot for adjectives	*black sheep*
OE	$[_{NP}$ D [A [N]]]	Slot for determiners	*the black sheep*
MdE	$[_{NP}$ M [D [A [N]]]]	Slot for peripheral mod.	*only the black sheep*

Crystallisation of noun phrase structure

- Slot for attributive modifiers: $[_{NP}$ A N]
  - Rise of specialised PoS:

    "The development of the adjective is perhaps one of the most conspicuous innovations in Germanic morphology. In Germanic the adjective is not only semantically deliminated by generally expressing some 'quality' (. . .), but it is also morphologically clearly definable." (Bammesberger 1992: 52-53)

  - Rise of contiguity: "'floating' adjectives become increasingly obsolete in Old Germanic languages (Van de Velde et al. 2014)

  - Rise of consistent prefield position: (non-relator, 'bare') postposition is diachronically decreasing in all Germanic languages

- Used increasingly: $[_{NP}$ A A ... N]
  - Increase in the occurrence of stacked adjectives (González-Díaz 2007 for English)

Discontinuous NPs in Old Germanic

- WEST GERMANIC

huand	*it*	*an*	***fastaro***	*nis*	***erthu***	*gitimbrid*
for	it	on	steady	not.is	ground	built

'for it is not built on steady ground' (Old Saxon, Van der Horst 2008:305)

- EAST GERMANIC

dauns	*sijum*	***woþi***
odor	we.are	sweet

'we are a sweet savour' (Gothic, Behaghel 1932:241)

- NORTH GERMANIC

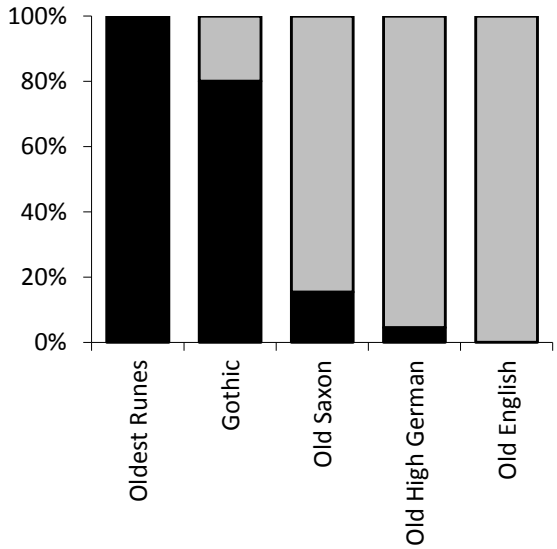
góðan	*eigum*	*vér*	***konung***
good	have	we	king

'we have a good king' (Old Norse, Faarlund 1994:56)

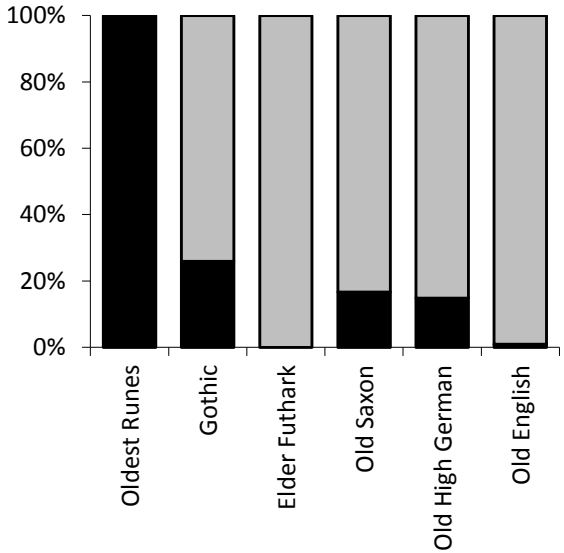
Increasing prefield position

Old Germanic languages
(Based on Smith 1971, see also Van de Velde 2009a)

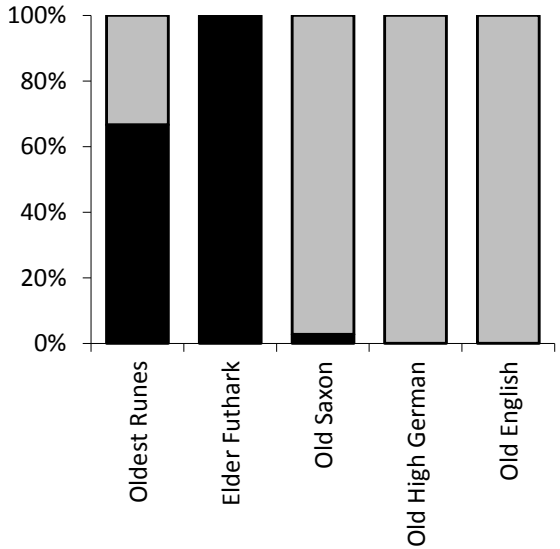
possessives



adjectives



demonstratives



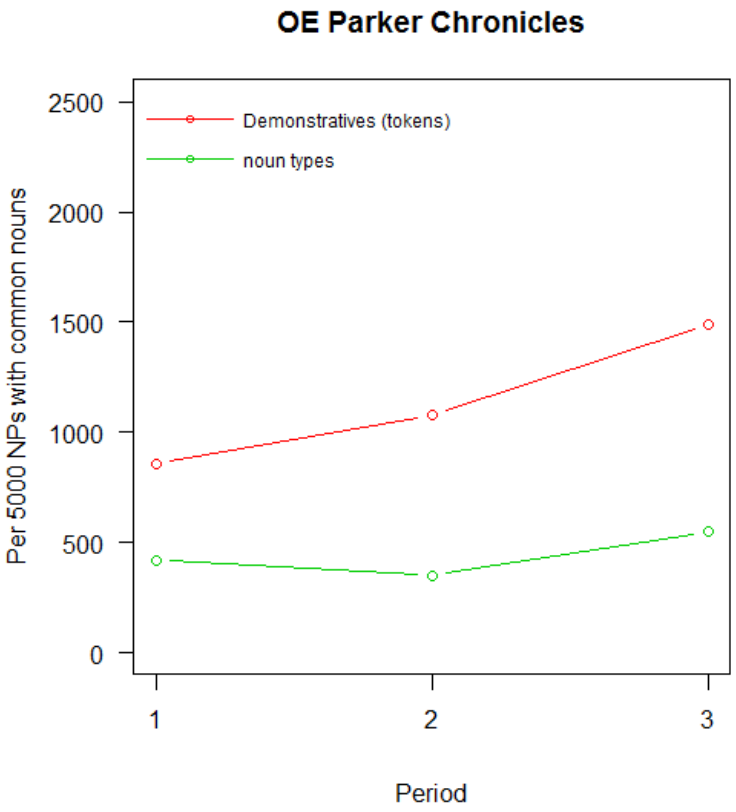
Crystallisation of noun phrase structure

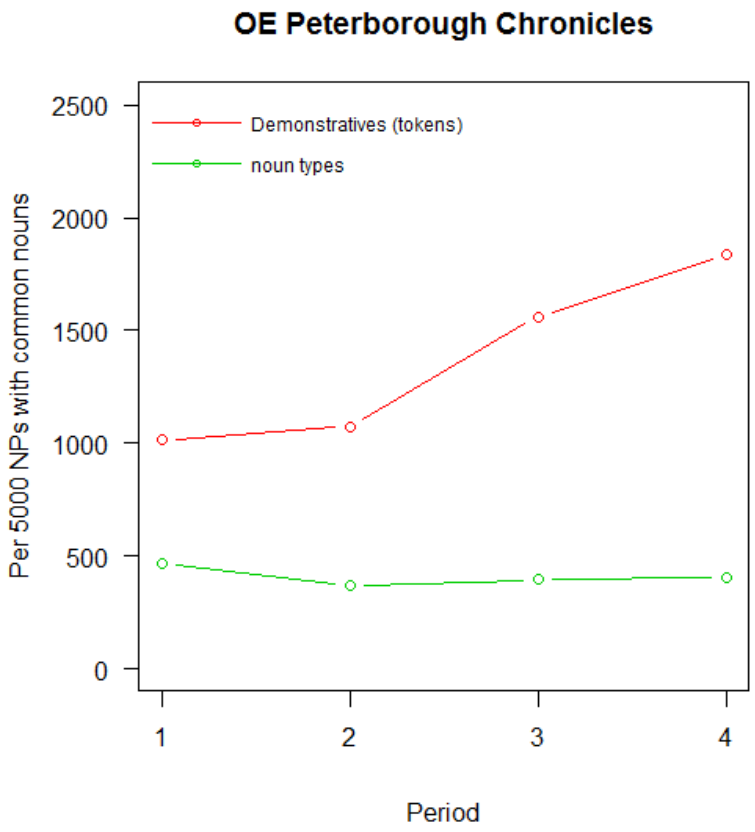
- Slot for attributive determiners: [$_{NP}$ D A N]
  - No article in PGm (Lehmann 1994:28)
  - By implication, no determiner slot (Himmelmann 1997: 133; Lyons 1999: 323)
  - Article emerges at the end of OE/OD/... (around 1000)

- Extension of the determiner: [$_{NP}$ D$_{Art\text{-}Dem,\ Poss\ ...}$ A N]
  - English: Spamer (1979); Roberts & Roussou (2003:ch.4); Denison (2006); Van de Velde (2009c), Sommerer (2011)
  - German: Plank (1992)
  - Dutch: Van de Velde (2010a, 2011b)

Increasing use of demonstratives

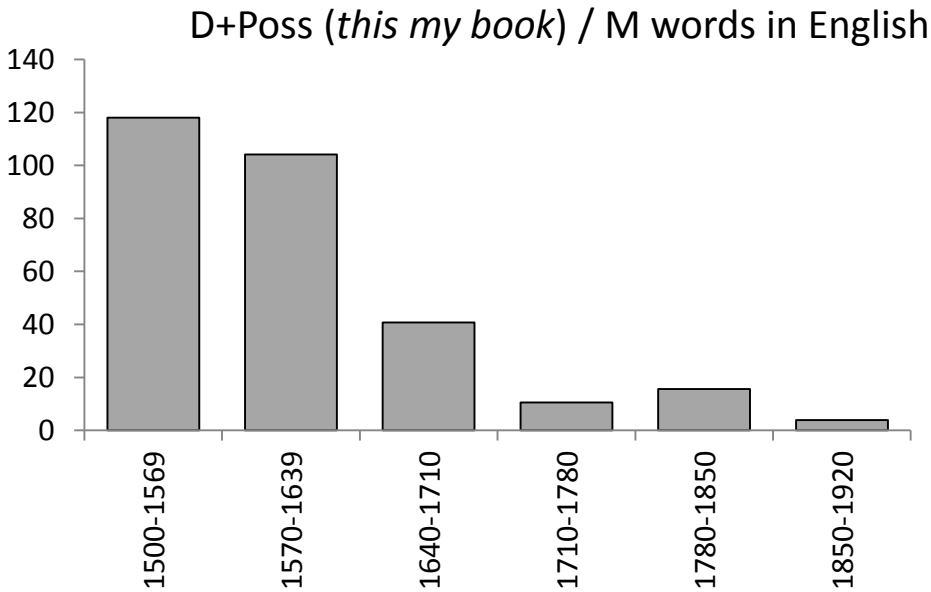
Old English
(Based on Sommerer 2011)





Crystallisation of noun phrase structure

- Extension of the determiner: [$_{NP}$ D$_{Art, Dem, Poss ...}$ A N] (Van de Velde 2009c)
  - One piece of evidence: applying criterion of mutual exclusivity (Dryer 2007: 161)



Incorporation of floating quantifiers

- [Q-CASE$_i$] … [$_{NP}$ … N-GENITIVE]]

  all *bagme* *godaize* (Gothic, Van de Velde 2009a:230)
  all:NOM.N.SG tree:GEN.M.PL good:GEN.M.PL
  'all good trees'

- [Q-CASE$_i$] … [$_{NP}$ … N-CASE$_i$]]

  alla *ufar* *insaht* (Gothic, Van de Velde 2009a:240)
  all:ACC.F.SG about explanation:ACC.F.SG
  'about every explanation'

- [Q-CASE$_i$] [$_{NP}$ … N-CASE$_i$]]

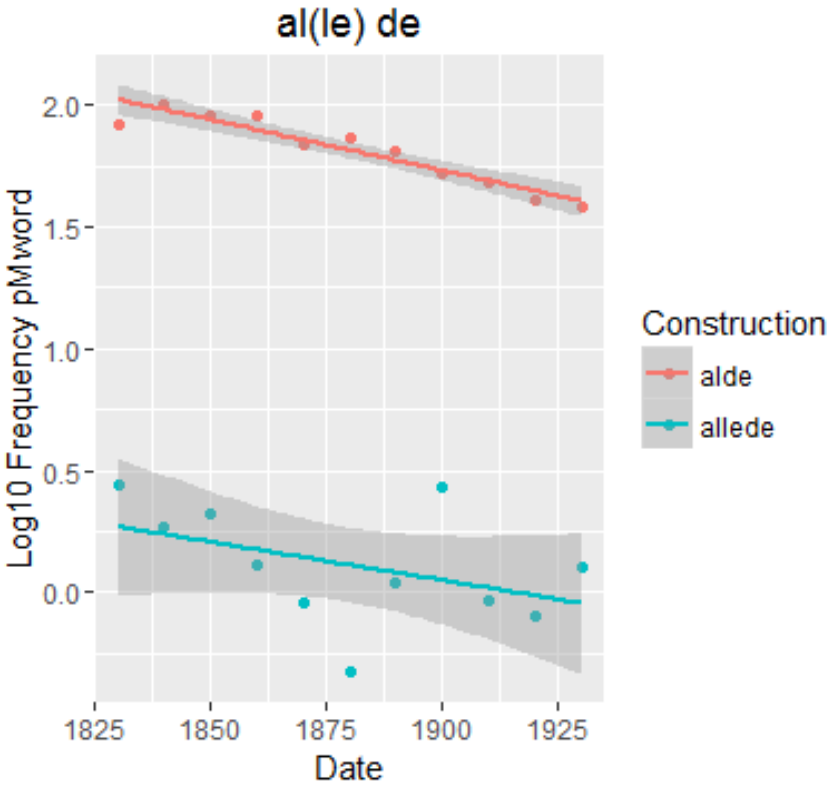
  allei *managein* (Gothic, Van de Velde 2009a:231)
  all:DAT.F.SG crowd:DAT.F.SG
  'the whole crowd'

Incorporation of floating quantifiers

Late Modern Dutch (Van de Velde 2015)

alle de mannen (all-INFL the men)	$[_{NP}$ Q] $[_{NP}$ D [N]]
al de mannen ('all-UNINFL the men')	[Q] $[_{NP}$ D [N]]
alle mannen (all-INFL men)	$[_{NP}$ Q N]

time



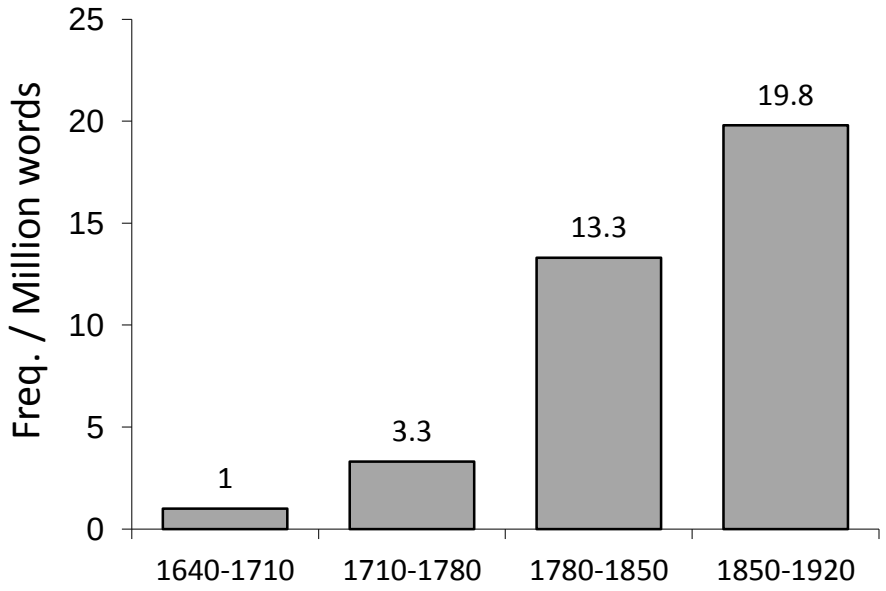
Crystallisation of noun phrase structure

- Slot for peripheral modifiers: [$_{NP}$ M D A N]
  - Rise of peripheral modifiers (Van de Velde 2010b, 2011a)
  - Starts out with focus particles (***even*** *the staff*), deriving from predicate adverbs

initial stage	*Þe barons portiond þe lond **euen** þam bituene* 'The barons divided the land equally between them.'	1330 (OED)
bridge context	*These sweet thoughts, doe **euen** refresh my labours*	1610 (OED)
switch context	*and concluded the horrid sport by kicking and mangling the heads, cutting of the lips, cheeks, ears, and noses; they **even** took out the jaw-bones, which they smokedried, together with the right hands, to carry home, as trophies of their victory*	1796 (OED)
form follows function	*The tone of insolent superiority assumed by **even** the gutter urchins.*	1863 (OED)

Crystallisation of noun phrase structure

- Extension of the peripheral modifiers: [$_{NP}$ $M_{FocP, SAdv ...}$ D A N]
  - Increase in use: PP-internal M ('with only the women') in English



  - Increase in complexity of M (Van de Velde 2011a)

Conclusions

- Emergence of phrasal structure
  - The change is NOT: 'rigidification of word order', as it is sometimes portrayed
  - The change is: emergence of NP structure as such, followed by gradual expansion
  - This gives a unified explanation for
    i. morphological changes (emergence of adjectival morphology)
    ii. successive steps in rigidification of word order (higher layers later)
    iii. lack of stacked modification in NP in former language stages
    iv. increasing use of complex NP modification (Leung & Van der Wurff 2009; Biber & Clark 2002)
    v. increasing reliance on NPs: Admoni (1967); Ebert (1978:49-50)
  - and does justice to the idea that phrase structure is not universal

Conclusions

- Emergence of NP structure: why?
  1. To dampen processing cost (Steels & Garcia Casademont 2015:44)
  2. Cultural changes (Perkins 1992; Lupyan & Dale 2010), including influence of written language (see Maas 2009 on German, Biber & Gray 2011 on the influence of writter registers, and Givón 2002:76 on the correlation between non-configurationality and spoken language): "[W]ritten language tends to have an 'integrated' quality which contrasts with the fragmented quality of spoken language." (Chafe 1982:38)
  3. L2-learners: esoteric > exoteric
- Explanation 1 is supported by the observation that the emergence of NP structure is not restricted to Europe (Heine & Kuteva 2007: 84-86, 279-287)
- Explanation 2 and 3 answer what in linguistics is known as the 'actuation problem'. Germanic is characterised by:
  - Great number of non-inherited vocabulary (Prokosch 1939)
  - Contact with other languages (Hawkins 1990, Roberge 2010 for an overview)
  - L2 acquisition continues to play a role in Early Modern period (Howell 2006, Lamiroy & Van de Velde 2010)